J. Fluid Mech. (2007), *vol.* 580, *pp.* 261–281. © 2007 Cambridge University Press
doi:10.1017/S0022112007005332 Printed in the United Kingdom

Optimal control of mixing in Stokes fluid flows

GEORGE MATHEW[1], IGOR MEZIĆ[1], SYMEON GRIVOPOULOS[1], UMESH VAIDYA[2] AND LINDA PETZOLD[1]

[1]Department of Mechanical Engineering, University of California, Santa Barbara, CA 93106, USA
[2]Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, USA

(Received 15 October 2005 and in revised form 7 December 2006)

Motivated by the problem of microfluidic mixing, optimal control of advective mixing in Stokes fluid flows is considered. The velocity field is assumed to be induced by a finite set of spatially distributed force fields that can be modulated arbitrarily with time, and a passive material is advected by the flow. To quantify the degree of mixedness of a density field, we use a Sobolev space norm of negative index. We frame a finite-time optimal control problem for which we aim to find the modulation that achieves the best mixing for a fixed value of the action (the time integral of the kinetic energy of the fluid body) per unit mass. We derive the first-order necessary conditions for optimality that can be expressed as a two-point boundary value problem (TPBVP) and discuss some elementary properties that the optimal controls must satisfy. A conjugate gradient descent method is used to solve the optimal control problem and we present numerical results for two problems involving arrays of vortices. A comparison of the mixing performance shows that optimal aperiodic inputs give better results than sinusoidal inputs with the same energy.

1. Introduction

From a practical engineering perspective, and in particular for microfluidic mixing devices (Stroock *et al.* 2002), it is often necessary to design protocols to mix a specific initial density field in finite time. The advent of new microfluidic technologies (see Stone, Stroock & Ajdari 2004) that make use of phenomena such as dielectrophoresis and electro-osmosis make it possible to generate arbitrarily time-varying velocity fields, thus opening the possibility of using the tools of optimal control theory to facilitate microfluidic mixing. The Péclet numbers (the ratio between advective and diffusive transport) for typical microscale flows are large ($Pe > 100$) and given the typical velocities at the microscale, the channel lengths required for diffusion alone to cause uniform spreading of material are prohibitively long ($\gg 1$ cm) (see Stroock *et al.* 2002). For this reason, microscale mixing processes need to be advection dominated.

Research on fluid mixing has been dominated by studies that are geometric in nature and applied to steady and time-periodic flows (see Aref 1984; Ottino 1989; Wiggins 1992). There have been relatively few studies of flows that are aperiodic in time (see e.g. Haller & Poje 1998; Malhotra & Wiggins 1998; Chrisohoides & Sotiropoulos 2003). In ergodic-theoretic approaches to mixing (Lasota & Mackey 1994; Arnold & Avez 1968; Petersen 1983), geometric properties of the underlying dynamical system are implicit and the emphasis is usually on infinite-time-averaged properties. An example of a study dealing with finite-time statistical properties associated with finite-time geometric properties is Poje, Haller & Mezić (1999). The objective of our paper is to

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

advance a new formalism for optimal control of advective mixing in aperiodic flows, by building upon the ideas for quantifying mixing that are discussed in our previous work (Mathew, Mezić & Petzold 2005) and combining it with concepts from finite-horizon optimal control theory. Previous approaches to the control of mixing include the study of control protocols that destroy symmetries (Franjione & Ottino 1992), maximization of mixing by enhancing Lyapunov exponents (Sharma & Gupte 1997), the study of optimal mixing protocols for combinations of shear flows (D'Alessandro, Dahleh & Mezić 1999), feedback control of mixing optimizing Eulerian-velocity-based measures of mixing (Balogh, Aamo & Kristic 2005), topological approaches (Boyland, Aref & Stremler 2000; Thiffeault & Finn 2006) and the control of unstable manifolds emanating from separation points (Wang *et al.* 2003). All these approaches aim to control geometric features of the flow and suffer from the drawback that the mixing protocols are designed without any consideration of the initial density field (or fluid configuration). Our approach in this paper is to set aside geometric aspects and to control the evolution of advected density fields in a manner that is optimal for mixing.

The problem considered in this paper is that of a Stokes fluid flow where the velocity field is assumed to be induced by a finite set of spatially distributed force fields that can be modulated arbitrarily in time. A passive material is advected by the flow. To quantify the degree of mixedness of a density field, we use a Sobolev space norm of negative index. The optimal control problem is to find the modulation of the force fields that achieves the best mixing for a fixed value of the action (the time integral of the kinetic energy of the fluid body) per unit mass. We derive the first-order necessary conditions for optimality that can be expressed as a two-point boundary value problem (TPBVP). We seek numerical approximations to solutions of the TPBVP using a conjugate gradient descent method. We also discuss two other related optimal control problems: where we require a fixed degree of mixedness to be achieved with minimum action; and where we minimize a cost-function that is a weighted sum of the degree of mixedness and the action per unit mass. A weighting parameter $w > 0$ determines how much the action is penalized relative to the degree of mixedness achieved. The relevance of the solutions that minimize this weighted sum is discussed in more detail in § 2.2.2.

In the planar case, it is impossible to achieve chaotic advection with a velocity field of the type $\alpha(t)\boldsymbol{u}(\boldsymbol{x})$, where $\boldsymbol{u}$ is a velocity field and α could be an arbitrary function of time. This is because the Lagrangian fluid elements are constrained to flow along the streamlines of the velocity field $\boldsymbol{u}$ and therefore it is impossible for a particle to visit every portion of the phase space. To cause chaotic advection, it is necessary to have at least two independent velocity fields that are superimposed upon each other and at least one is modulated with time. In this paper, we assume velocity fields of the form $\alpha_1(t)\boldsymbol{u}_1(\boldsymbol{x}) + \alpha_2(t)\boldsymbol{u}_2(\boldsymbol{x})$. In the three-dimensional case, it is possible to have chaotic advection even with a steady velocity field (Bajer & Moffatt 1990; Stone, Nadim & Strogatz 1991, Fountain *et al.* 2000). However, in the presence of symmetries (Mezić & Wiggins 1994; Grigoriev 2005) it may be necessary to have two or three independent velocity fields with complex time-dependence to cause efficient mixing. For simplicity, all the discussion and examples in this paper are for two-dimensional toroidal domains, i.e. square domains with periodic boundary conditions in both coordinates.

2. Description of the problem

2.1. *Evolution equations*

Consider the incompressible Navier–Stokes equation with external body forcing, with the appropriate scalings for a Stokes flow regime on a two-dimensional toroidal

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

domain $T^2 = [0, 1]^2$:

$$(Re\,St)\boldsymbol{u}_t + (Re)\boldsymbol{u}\cdot\nabla\boldsymbol{u} = -\nabla p + \boldsymbol{F} + \nabla^2\boldsymbol{u}, \quad \nabla\cdot\boldsymbol{u} = 0, \tag{2.1}$$

where $Re = \rho UL/\mu$ is the Reynolds number, $St = L/UT$ the Strouhal number, U a characteristic velocity, L a length scale, and T a time scale. All vectors are written in bold font (eg: $\boldsymbol{x}$) and their respective elements are written in usual font with indices as subscripts (e.g. $x_1, x_2, \ldots$). Subscript t (e.g. c_t) indicates partial differentiation with time. ∇ is the gradient. The density of the fluid is ρ and its viscosity μ. For $Re \to 0$ and $St \to 0$, the equation reduces to the Stokes equation

$$\nabla^2\boldsymbol{u} = \nabla p - \boldsymbol{F}, \quad \nabla\cdot\boldsymbol{u} = 0. \tag{2.2}$$

We assume the force field to be of the form $\boldsymbol{F}(\boldsymbol{x}, t) = \sum_{i=1}^{n} \alpha_i(t)\boldsymbol{F}_i(\boldsymbol{x})$. If $\boldsymbol{u}_i$ is the velocity field induced by the force field $\boldsymbol{F}_i$ (i.e. $\boldsymbol{u}_i$ is the solution to equation (2.2) for $\boldsymbol{F} = \boldsymbol{F}_i$), then the velocity field can be written as

$$\boldsymbol{u}(\boldsymbol{x}, t) = \sum_{i=1}^{n} \alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x}). \tag{2.3}$$

If it were possible to create arbitrary force fields, we could assume the velocity field to be of the form (2.3), but with possibly infinite modes chosen from a convenient basis, and proceed to consider scalar field control as in what follows, and then return to the Stokes equation or Navier–Stokes equation to find the appropriate force $\boldsymbol{F}(\boldsymbol{x}, t)$ to generate $\boldsymbol{u}(\boldsymbol{x}, t)$. But, this is not likely to be realizable in practice. Here, we study the special case where it is possible to generate a finite set of force fields ($\boldsymbol{F}_i$) using some mechanisms, and assume that they can be modulated in time arbitrarily. Then we find the corresponding $\boldsymbol{u}_i$ and find the modulation in time that is optimal for mixing. Now, consider a density field $c(\boldsymbol{x}, t)$ that is being advected by the velocity field (2.3). Then

$$c_t(\boldsymbol{x}, t) = -\boldsymbol{u}(\boldsymbol{x}, t)\cdot\nabla c(\boldsymbol{x}, t) = -\sum_{i=1}^{n} \alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c(\boldsymbol{x}, t), \quad c(\boldsymbol{x}, 0) = c_0(\boldsymbol{x}). \tag{2.4}$$

The control system (2.4) is a bilinear system where $c(., t)$ is the infinite-dimensional state and $\boldsymbol{\alpha}(t)$ is the finite-dimensional control input. Issues of stability and controllability have been studied for both finite-dimensional bilinear systems (see e.g. Jurdjevic & Quinn 1978; Slemrod 1978) and infinite-dimensional bilinear systems (see e.g. Ball, Marsden & Slemrod 1982).

Note that since each $\boldsymbol{u}_i$ is volume-preserving and there is no diffusion, the mean and variance of the density field is preserved under the evolution by equation (2.4). In fluid-mechanical terms, this corresponds to the limit as the Péclet number (Pe) goes to infinity. Recall that the Péclet number is the ratio of advective transport to diffusive transport. In any realistic system, there would be some non-zero diffusion. But in a typical mixing process there is an initial stage during which the diffusive effects are negligible and the variance remains almost constant. It is during this initial stage, which is commonly referred to as 'stirring', that stretching and folding of fluid elements occur and eventually facilitate diffusion to efficiently homogenize the passive material. Therefore, to optimize this initial stage of stirring, we need to consider the pure advection problem and use a measure for mixing different from the variance.

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

2.2. *The cost function and the optimal control problem*

2.2.1. *Measure of mixedness*

In our previous work (Mathew *et al.* 2005), a multiscale measure for mixing, referred to as the mix-norm, was presented to quantify the degree of mixedness of a density field. This measure is based on the concept of weak convergence and is related to the classical ergodic-theoretic notion of mixing. The definition of the measure is based on averaging the function over all scales and integrating the L^2 norms of the averaged functions over all scales. On an m-dimensional toroidal domain, the mix-norm $\Phi(c)$ is defined as follows:

$$\text{For } s \in (0,1),\ \phi^2(c,s) = \int_{T^m} d^2(c,\boldsymbol{p},s)\,\mathrm{d}\boldsymbol{p}, \quad \text{where } d(c,\boldsymbol{p},s) = \frac{\displaystyle\int_{B(\boldsymbol{p},s)} c(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}}{V(s)}. \tag{2.5}$$

$V(s)$ is the volume of the spherical set $B(\boldsymbol{p},s) = \{\boldsymbol{y} : \|\boldsymbol{y}-\boldsymbol{p}\| \leqslant \frac{1}{2}s\}$, $d(c,\boldsymbol{p},s)$ is the average value of the function c within the spherical set $B(\boldsymbol{p},s)$ and $\phi(c,s)$ is the L^2 norm of the averaged function $d(c,.,s)$ at a fixed scale s. Then $\Phi(c)$ is defined as

$$\Phi^2(c) = \int_0^1 \phi^2(c,s)\,\mathrm{d}s. \tag{2.6}$$

$\Phi^2(c)$ can also be written as the inner product $\langle c, [M]c\rangle$, where $\langle .,.\rangle$ is the standard inner product and $[M]$ is a self-adjoint, positive semi-definite and spatially invariant operator. Thus its eigenfunctions are the Fourier basis functions (i.e. $\mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{k}\cdot\boldsymbol{x})}$). The eigenvalue $\boldsymbol{\Lambda}_{\boldsymbol{k}}$ corresponding to a Fourier mode with wavenumber vector $\boldsymbol{k}$ satisfies the inequality

$$\frac{\mu_1}{(1+(2\pi\|\boldsymbol{k}\|)^2)^{1/2}} \leqslant \boldsymbol{\Lambda}_{\boldsymbol{k}} \leqslant \frac{\mu_2}{(1+(2\pi\|\boldsymbol{k}\|)^2)^{1/2}}, \tag{2.7}$$

where $\mu_1, \mu_2 > 0$ are constants. For a density field with a Fourier expansion $c(\boldsymbol{x}) = \sum_{\boldsymbol{k}} c_{\boldsymbol{k}}\,\mathrm{e}^{\mathrm{i}2\pi(\boldsymbol{k}\cdot\boldsymbol{x})}$, the mix-norm can be computed as $\Phi^2(c) = \sum_{\boldsymbol{k}} \Lambda_{\boldsymbol{k}}|c_{\boldsymbol{k}}|^2$. It follows from inequality (2.7) that the mix-norm is equivalent to a Sobolev space norm of negative index $s = -\frac{1}{2}$, which is defined for every $c \in L^2_{T^m}$ as

$$\|c\|_{H^{-1/2}} = \left(\sum_{\boldsymbol{k}} \frac{1}{(1+(2\pi\|\boldsymbol{k}\|)^2)^{1/2}}|c_{\boldsymbol{k}}|^2\right)^{1/2}, \tag{2.8}$$

i.e. for all $c \in L^2_{T^m}$,

$$\sqrt{\mu_1}\|c\|_{H^{-1/2}} \leqslant \Phi(c) \leqslant \sqrt{\mu_2}\|c\|_{H^{-1/2}}. \tag{2.9}$$

The mathematical equivalence of the mix-norm and the $\|\cdot\|_{H^{-1/2}}$ norm makes them interchangeable. If one of the norms decays to zero for a sequence of functions, it implies the decay of the other. Moreover, the asymptotic rate of decay of both norms would be exactly the same. In particular, if the $\|\cdot\|_{H^{-1/2}}$ norm can be shown to decay exponentially to zero for a given sequence of functions, it implies that the magnitudes of average function values over almost every spherical set converge to zero at an exponential rate. In all further references and computations of the mix-norm, we use the $\|\cdot\|_{H^{-1/2}}$ norm. In other words, the linear operator $[M]$ is assumed to have eigenvalues $\Lambda_{\boldsymbol{k}} = (1+(2\pi\|\boldsymbol{k}\|)^2)^{-1/2}$. Note that when computing $\Phi^2(c)$, the eigenvalues $\Lambda_{\boldsymbol{k}}$ act as a weighting on the energy contained in various Fourier modes. The larger the magnitude of the wavenumber vector, the smaller the weighting. In

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

all our computations, we use FFT software (Frigo & Johnson 1998) to compute the mix-norm.

The mix-variance of a density field c is defined as $\Phi^2(c-\bar{c})$, where $\bar{c}$ is the mean of the field c over the whole space. It can be proven that the underlying flow field is mixing in the classical ergodic-theoretic sense if and only if the mix-variance of any advected density field decays to zero or equivalently the evolving density field converges weakly to a field of constant value $\bar{c}$ (see Mathew *et al.* 2005). Decay of the mix-variance of an evolving density field implies that energy is transferred from low-wavenumber Fourier modes to high-wavenumber Fourier modes. In what follows, we will always assume that c has zero mean and therefore the mix-variance would be the same as the square of the mix-norm. For the case of a density field c with non-zero mean, one can define the new field $d = c - \bar{c}$ and study the evolution of d.

A measure for mixing that is similar in spirit to the mix-variance is discussed by Stone & Stone (2005), where the mixing measure used is a scale-independent quantity and is based on how rapidly diffusion alone would homogenize a density field. In fact, a generalization of the measure introduced by Stone & Stone (2005) may be obtained so that it captures the mixing properties that the mix-variance does. The exact relationship between these two quantities needs to be investigated further.

2.2.2. *The optimal control problem*

In optimal control problems, it is common to have various competing performance objectives. Here, we use two such objectives. One is the degree of mixedness of the density field at the final time, which is quantified in terms of the mix-variance, and the second is the action, which is the time integral of the kinetic energy per unit mass of the fluid body. (Throughout the rest of the paper, the term 'action' actually refers to the action per unit mass.) Computing the kinetic energy per unit mass at time t, we have

$$\begin{aligned}\frac{\frac{1}{2}\displaystyle\int_{T^2}\rho\boldsymbol{u}(\boldsymbol{x},t)\cdot\boldsymbol{u}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}}{\displaystyle\int_{T^2}\rho\,\mathrm{d}\boldsymbol{x}} &= \frac{1}{2}\int_{T^2}\left(\sum_{i=1}^{n}\alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\right)\cdot\left(\sum_{i=1}^{n}\alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\right)\mathrm{d}\boldsymbol{x}\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i(t)\left(\frac{1}{2}\int_{T^2}\boldsymbol{u}_i(\boldsymbol{x})\cdot\boldsymbol{u}_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)\alpha_j(t)\\ &= \boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}(t),\end{aligned} \tag{2.10}$$

where

$$\boldsymbol{R}_{i,j} := \frac{1}{2}\int_{T^2}\boldsymbol{u}_i(\boldsymbol{x})\cdot\boldsymbol{u}_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Then the action caused by a control $\boldsymbol{\alpha}$ is given as

$$A(\boldsymbol{\alpha}) := \int_0^{t_f}\boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}(t)\,\mathrm{d}t. \tag{2.11}$$

In the rest of the paper, we will assume that the set of velocity fields $\{\boldsymbol{u}_i\}$ are linearly independent. This guarantees that the matrix R will be symmetric positive definite. Given below are three possible optimal control problems.

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

Optimal Control Problem I: Fixed action caused by the controls:

$$\begin{aligned} &\underset{\boldsymbol{\alpha}\in L^2_{[0,t_f]}}{\text{Min}} \{W(\boldsymbol{\alpha}) := \langle c(.,t_f), [M]c(.,t_f)\rangle\} \\ &\text{subject to (2.4) and such that } A(\boldsymbol{\alpha}) = A^*. \end{aligned} \tag{2.12}$$

Optimal Control Problem II: Fixed degree of mixedness at the final time:

$$\begin{aligned} &\underset{\boldsymbol{\alpha}\in L^2_{[0,t_f]}}{\text{Min}} \left\{ A(\boldsymbol{\alpha}) := \int_0^{t_f} \boldsymbol{\alpha}(t)\cdot \boldsymbol{R}\boldsymbol{\alpha}(t)\,\mathrm{d}t \right\} \\ &\text{subject to (2.4) and such that } \langle c(.,t_f), [M]c(.,t_f)\rangle = (\Phi^*)^2. \end{aligned} \tag{2.13}$$

Optimal Control Problem III: Weighted combination of mixedness and action as the cost-function:

$$\begin{aligned} &\underset{\boldsymbol{\alpha}\in L^2_{[0,t_f]}}{\text{Min}} \left\{ C(\boldsymbol{\alpha}) := \langle c(.,t_f), [M]c(.,t_f)\rangle + w\cdot \int_0^{t_f} \boldsymbol{\alpha}(t)\cdot \boldsymbol{R}\boldsymbol{\alpha}(t)\,\mathrm{d}t \right\} \\ &\text{subject to (2.4).} \end{aligned} \tag{2.14}$$

In these problems t_f is a finite time-span and $w>0$ is a weighting parameter that determines how much the action caused by the control should be penalized relative to the mixedness. In general, the matrix $\boldsymbol{R}$ may be set arbitrarily, but here we choose $\boldsymbol{R}$ as discussed above because of the nice physical interpretation of the resulting cost-function. Solutions to the optimal control problem (2.14) are Pareto optimal solutions in the following sense. Let $\boldsymbol{\alpha}^*$ be the $\boldsymbol{\alpha}$ that minimizes $C(\boldsymbol{\alpha})$ for some w. Let $\boldsymbol{\alpha}^*$ lead to a certain degree of mixedness $(\Phi^*)^2$ and to a certain value for the action A^*. Then $\boldsymbol{\alpha}^*$ is a solution to both optimal control problems I and II.

The argument that $\boldsymbol{\alpha}^*$ also solves optimal control problem II is as follows. The proof is by contradiction. Let $\bar{\boldsymbol{\alpha}}$ be a solution to optimal control problem II. Therefore, $\bar{\boldsymbol{\alpha}}$ achieves a degree of mixedness $(\Phi^*)^2$. Now, assume that $A(\bar{\boldsymbol{\alpha}})<A^*$. Then $C(\bar{\boldsymbol{\alpha}})<C(\boldsymbol{\alpha}^*)=(\Phi^*)^2+wA^*$, and therefore $\bar{\boldsymbol{\alpha}}$ would be a better solution to optimal control problem III than $\boldsymbol{\alpha}^*$, thus contradicting the assumption that $\boldsymbol{\alpha}^*$ minimizes $C(\boldsymbol{\alpha})$. Now, assume $\bar{\boldsymbol{\alpha}}$ is such that $A(\bar{\boldsymbol{\alpha}})>A^*$. But, then $A(\boldsymbol{\alpha}^*)<A(\bar{\boldsymbol{\alpha}})$, thus contradicting the assumption that $\bar{\boldsymbol{\alpha}}$ is a solution to optimal control problem II. Therefore, it is necessary that $A(\boldsymbol{\alpha}^*)=A(\bar{\boldsymbol{\alpha}})$. A similar argument shows that $\boldsymbol{\alpha}^*$ also solves optimal control problem I.

The existence of solutions to optimal control problem II depends on the issue of controllability i.e. it must be possible to achieve the desired degree of mixedness with the available velocity fields, whereas the existence of solutions to optimal control problems I and III are irrespective of this issue of controllability. The relevance of each of these optimal control problems may depend on the specific engineering situation. If one is interested only in finding Pareto optimal solutions, one could do so by generating solutions to optimal control problem III for various values of w. In this paper, for convenience, we focus on optimal control problem I, i.e. we aim to achieve optimal mixing for a fixed value of the action.

3. Solving the optimal control problem

Solving the optimal control problems discussed in the previous section is both analytically and numerically challenging. There are only limited results available for optimal control of infinite-dimensional bilinear systems (see e.g. Banks 1987). Here,

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

we use a gradient-based method to find numerical approximations of solutions to the optimal control problem. The partial differential equation (2.4) is treated as a constraint and using the Lagrange multiplier formalism, an augmented functional $\widehat{W}$ is defined as

$$\widehat{W}(\boldsymbol{\alpha}, c, \boldsymbol{\lambda}, z) := \langle c(.,t_f), [M]c(.,t_f)\rangle - z.\left(A^* - \int_0^{t_f} \boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}(t)\,\mathrm{d}t\right) - \int_0^{t_f}\int_{T^2} \boldsymbol{\lambda}(\boldsymbol{x},t)\left(c_t(\boldsymbol{x},t) + \sum_{i=1}^{n}\alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \quad (3.1)$$

The variables z and $\boldsymbol{\lambda}(\boldsymbol{x},t)$ play the role of Lagrange multipliers. Finding the Fréchet derivatives of $\widehat{W}$ with respect to $\boldsymbol{\alpha}$, c, $\boldsymbol{\lambda}$ and z and setting them to zero gives the first-order necessary conditions for optimality in terms of the two-point boundary value problem (TPBVP):

$$\left[\frac{\mathrm{D}\widehat{W}}{\mathrm{D}\boldsymbol{\lambda}}\right]\partial\boldsymbol{\lambda} = 0; \Longrightarrow c_t^*(\boldsymbol{x},t) + \sum_{i=1}^{n}\alpha_i^*(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c^*(\boldsymbol{x},t) = 0, \quad c^*(\boldsymbol{x},0) = c_0(\boldsymbol{x}), \quad (3.2a)$$

$$\left[\frac{\mathrm{D}\widehat{W}}{\mathrm{D}c}\right]\partial c = 0; \Longrightarrow \boldsymbol{\lambda}_t^*(\boldsymbol{x},t) + \sum_{i=1}^{n}\alpha_i^*(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla\boldsymbol{\lambda}^*(\boldsymbol{x},t) = 0, \quad \boldsymbol{\lambda}^*(\boldsymbol{x},t_f) = 2[M]c^*(\boldsymbol{x},t_f), \quad (3.2b)$$

$$\left[\frac{\mathrm{D}\widehat{W}}{\mathrm{D}z}\right]\partial z = 0; \Longrightarrow \int_0^{t_f}\boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}(t)\,\mathrm{d}t = A^*, \quad (3.2c)$$

$$\left[\frac{\mathrm{D}\widehat{W}}{\mathrm{D}\boldsymbol{\alpha}}\right]\boldsymbol{\partial\alpha} = 0; \Longrightarrow 2z\boldsymbol{R}\boldsymbol{\alpha}^*(t) + \boldsymbol{\beta}^*(t) = 0, \quad (3.2d)$$

where

$$\boldsymbol{\beta}_i^*(t) = -\langle\boldsymbol{\lambda}^*(.,t), \boldsymbol{u}_i(\cdot)\cdot\nabla c^*(.,t)\rangle \quad \text{for } i = 1, 2, \ldots, n. \quad (3.2e)$$

For the basic concepts underlying the derivation of these necessary conditions, see Gelfand & Fomin (1963) and Kirk (1970). For completeness, we give the derivation of the TPBVP (3.2) in the Appendix. The variable $\boldsymbol{\lambda}(x,t)$ would be referred to as the costate field. $\boldsymbol{\alpha}^*$ is the optimal control and c^* and $\boldsymbol{\lambda}^*$ are the solutions of the state and costate fields (c and $\boldsymbol{\lambda}$) corresponding to the optimal control.

3.1. *Conservation of kinetic energy by the optimal controls*

An important feature of the optimal controls is that the resulting system evolves in a manner such that the kinetic energy of the fluid body is conserved, i.e. the solutions $\boldsymbol{\alpha}^*$ to the TPBVP are such that

$$\boldsymbol{\alpha}^*(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}^*(t) = \text{constant}. \quad (3.3)$$

This statement can be verified as follows. From equation (3.2) we have

$$\boldsymbol{\alpha}^*(t) = \frac{-1}{2z}\boldsymbol{R}^{-1}\boldsymbol{\beta}^*(t). \quad (3.4)$$

Computing the time-derivative of $\boldsymbol{\alpha}^*(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}^*(t)$, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{\alpha}^*(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}^*(t)) = 2\boldsymbol{\alpha}^*(t)\cdot\boldsymbol{R}\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{\alpha}^*(t)) = 2\boldsymbol{\alpha}^*(t)\cdot\boldsymbol{R}\left(\frac{-1}{2z}\boldsymbol{R}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{\beta}^*(t))\right)$$
$$= \frac{-1}{z}\boldsymbol{\alpha}^*(t)\cdot\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{\beta}^*(t)) = \frac{-1}{z}\sum_{i=1}^{n}\alpha_i^*(t)\frac{\mathrm{d}\beta_i^*(t)}{\mathrm{d}t}. \quad (3.5)$$

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

Now,

$$\begin{aligned}\frac{\mathrm{d}\beta_i^*(t)}{\mathrm{d}t} &= -\langle \lambda_t^*(.,t), \boldsymbol{u}_i(.)\cdot\nabla c^*(.,t)\rangle + \langle \boldsymbol{u}_i(\cdot)\cdot\nabla\boldsymbol{\lambda}^*(.,t), c_t^*(.,t)\rangle \\ &= \langle \boldsymbol{u}(.,t)\cdot\nabla\boldsymbol{\lambda}^*(.,t), \boldsymbol{u}_i(\cdot)\cdot\nabla c^*(.,t)\rangle - \langle \boldsymbol{u}_i(\cdot)\cdot\nabla\boldsymbol{\lambda}^*(.,t), \boldsymbol{u}(.,t)\cdot\nabla c^*(.,t)\rangle. \end{aligned} \tag{3.6}$$

Therefore,

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\alpha}^*(t)\cdot R\boldsymbol{\alpha}^*(t)\right) \\ &= \frac{-1}{z}\sum_{i=1}^{n}\alpha_i^*(t)(\langle \boldsymbol{u}(.,t)\cdot\nabla\boldsymbol{\lambda}^*(.,t), \boldsymbol{u}_i(\cdot)\cdot\nabla c^*(.,t)\rangle - \langle \boldsymbol{u}_i(\cdot)\cdot\nabla\boldsymbol{\lambda}^*(.,t), \boldsymbol{u}(.,t)\cdot\nabla c^*(.,t)\rangle) \\ &= \frac{-1}{z}\left\langle \boldsymbol{u}(.,t)\cdot\nabla\boldsymbol{\lambda}^*(.,t), \sum_{i=1}^{n}\alpha_i^*(t)\boldsymbol{u}_i(\cdot)\cdot\nabla c^*(.,t)\right\rangle \\ &\quad + \frac{1}{z}\left\langle \sum_{i=1}^{n}\alpha_i^*(t)\boldsymbol{u}_i(\cdot)\cdot\nabla\boldsymbol{\lambda}^*(.,t), \boldsymbol{u}(.,t)\cdot\nabla c^*(.,t)\right\rangle \\ &= \frac{-1}{z}\langle \boldsymbol{u}(.,t)\cdot\nabla\boldsymbol{\lambda}^*(.,t), \boldsymbol{u}(.,t)\cdot\nabla c^*(.,t)\rangle + \frac{1}{z}\langle \boldsymbol{u}(.,t)\cdot\nabla\boldsymbol{\lambda}^*(.,t), \boldsymbol{u}(.,t)\cdot\nabla c^*(.,t)\rangle = 0. \end{aligned} \tag{3.7}$$

Conservation of kinetic energy implies that the vector $\boldsymbol{\alpha}^*(t)$ evolves on the ellipsoid described by equation (3.3). For the special case when there are just two basis velocity fields (two controls) and when $\boldsymbol{R}$ is a diagonal matrix, the following properties hold for all the optimal controls:

(i) if the magnitude of one of the control inputs is increasing, the magnitude of the other control input must be decreasing;

(ii) when the time-derivative of one of the control inputs is zero, the other control input must either have zero time-derivative or actual value zero, i.e. as long as both control inputs are away from zero, maximum or minimum peaks of one of the control inputs must appear simultaneously with the maximum or minimum peaks of the other.

In the next section, we describe how we find numerical approximations of the solutions to the two-point boundary value problem using a conjugate gradient descent method (Luenberger 1984).

3.2. *Description of the conjugate gradient descent method*

For a given $\boldsymbol{\alpha}$, solutions of the density field c and costate field $\boldsymbol{\lambda}$ are explicitly known in terms of Lagrangian tracer particle trajectories. For this purpose we need to introduce the mapping $S_{\boldsymbol{\alpha}}$ that is defined as follows. For $t_f \geqslant t_0$, $S_{\boldsymbol{\alpha}}$ is such that if

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \sum_{i=1}^{n}\alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x}(t)), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0, \tag{3.8}$$

then $\boldsymbol{x}(t_f) = S_{\boldsymbol{\alpha}}(\boldsymbol{x}_0, t_0, t_f)$. Also, for $t_f < t_0$, $S_{\boldsymbol{\alpha}}$ is such that $S_{\boldsymbol{\alpha}}(., t_0, t_f) = S_{\boldsymbol{\alpha}}^{-1}(., t_f, t_0)$. Therefore, for a given $\boldsymbol{\alpha}$, the state and costate fields can be computed as

$$c(\boldsymbol{x}, t) = c_0(S_{\boldsymbol{\alpha}}(\boldsymbol{x}, t, 0)), \qquad \boldsymbol{\lambda}(\boldsymbol{x}, t) = \boldsymbol{\lambda}(S_{\boldsymbol{\alpha}}(\boldsymbol{x}, t, t_f), t_f) = 2[M]c(S_{\boldsymbol{\alpha}}(\boldsymbol{x}, t, t_f), t_f). \tag{3.9}$$

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

Note that, at a given $\boldsymbol{\alpha}$, the Fréchet derivative of the cost-function $W = \langle c(., t_f), [M]c(., t_f)\rangle$ with respect to $\boldsymbol{\alpha}$ is given as

$$\left[\frac{\mathrm{D}W}{\mathrm{D}\boldsymbol{\alpha}}\right]\partial\boldsymbol{\alpha} = \int_0^{t_f} \boldsymbol{\beta}(t)\cdot\partial\boldsymbol{\alpha}(t)\,\mathrm{d}t, \tag{3.10}$$

where

$$\beta_i(t) := -\langle \boldsymbol{\lambda}(., t), \boldsymbol{u}_i(\cdot)\cdot\nabla c(., t)\rangle.$$

In equations (3.9) and (3.10), the implicit dependence of the variables $\boldsymbol{\beta}$, c and $\boldsymbol{\lambda}$ on $\boldsymbol{\alpha}$ is suppressed in the notation. It is assumed that c and $\boldsymbol{\lambda}$ are the solutions of the respective partial differential equations corresponding to the $\boldsymbol{\alpha}$ under consideration. As described in the previous section, the optimal controls are such that the kinetic energy of the fluid body is conserved or the vector $\boldsymbol{\alpha}^*(t)$ evolves on the ellipsoid described by $\boldsymbol{\alpha}^*(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}^*(t) = \text{constant}$. Therefore, it is desirable to search for solutions that satisfy this property. For this purpose, we describe how we find approximations to the optimal controls for the case when there are two basis velocity fields and such that the kinetic energy is conserved. We assume the optimal controls to be of the form

$$\boldsymbol{\alpha}(t) = \boldsymbol{R}^{-1/2}\boldsymbol{b}(t), \tag{3.11}$$

where

$$b_1(t) = a\cos(\theta(t)), \quad b_2(t) = a\sin(\theta(t)). \tag{3.12}$$

This ensures that $\boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}(t) = a^2$ is a constant. (Note that $b_2(t) = -a\sin(\theta(t))$ is also a possibility.) Also, since the action is fixed at A^*, we have $a = \sqrt{A^*/t_f}$. Thus, we only need to optimize for the angle function θ to minimize the cost-function W. The Fréchet derivative of the cost-function W with respect to θ is given as

$$\left[\frac{\mathrm{D}W}{\mathrm{D}\,\theta}\right]\partial\theta = \left[\frac{\mathrm{D}W}{\mathrm{D}\boldsymbol{\alpha}}\right]\circ\left[\frac{\mathrm{D}\boldsymbol{\alpha}}{\mathrm{D}\theta}\right]\partial\theta = \int_0^{t_f}\partial W_\theta(t)\partial\theta(t)\,\mathrm{d}t, \tag{3.13}$$

where

$$\partial W_\theta(t) := \boldsymbol{\beta}(t)\cdot\boldsymbol{R}^{-1/2}\frac{\partial(\boldsymbol{b}(t))}{\partial\theta}.$$

Now, an initial guess $(\theta)^0$ is made for θ. The corresponding solutions for the state and costate fields, $(c)^0$ and $(\boldsymbol{\lambda})^0$, are computed using equation (3.9). Once $(c)^0$ and $(\boldsymbol{\lambda})^0$ are computed, the following iteration is performed:

$$\left.\begin{aligned}(\theta)^{n+1}(t) &= (\theta)^n(t) + (h)^n(s)^n(t),\\ (s)^{n+1}(t) &= -(\partial W_\theta)^{n+1}(t) + (\gamma)^n(s)^n(t),\end{aligned}\right\} \tag{3.14}$$

where $(\partial W_\theta)^{n+1}$ is ∂W_θ computed for $\theta = (\theta)^{n+1}$; $(h)^n$ is a scalar chosen so that it minimizes $W((\theta)^n + h(s)^n)$, i.e.

$$W((\theta)^n + (h)^n(s)^n) = \underset{h>0}{\mathrm{Min}}\, W((\theta)^n + h(s)^n). \tag{3.15}$$

The initial search direction $(s)^0$ is set to $(\partial W_\theta)^0$; $(\gamma)^n$ is a scalar computed using the so-called Polak–Ribiere formula

$$(\gamma)^n = \max\left\{\frac{\langle(\partial W_\theta)^{n+1}, (\partial W_\theta)^{n+1} - (\partial W_\theta)^n\rangle}{\langle(\partial W_\theta)^n, (\partial W_\theta)^n\rangle}, 0\right\}. \tag{3.16}$$

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

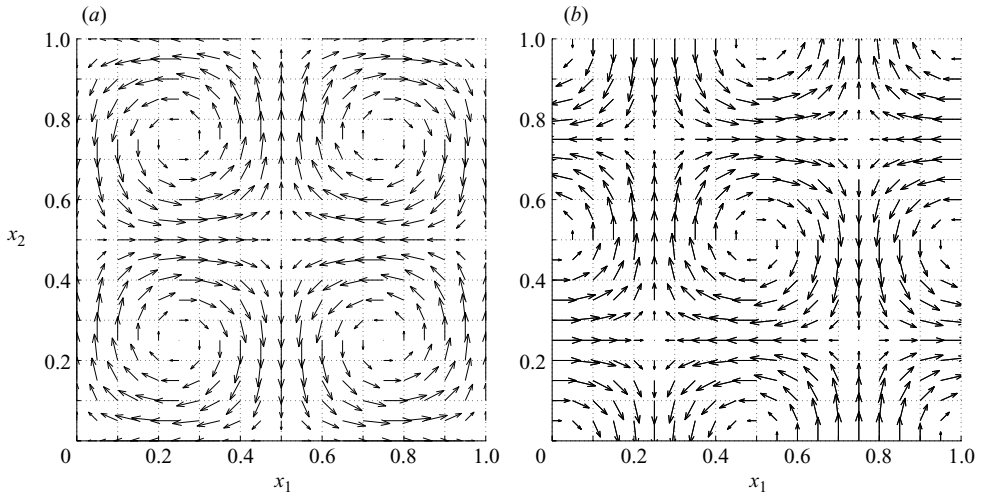


FIGURE 1. Velocity fields defined in (4.1): (*a*) $\boldsymbol{u}_1$, (*b*) $\boldsymbol{u}_2$.

When $(\gamma)^n$ is identically set to zero, the method reduces to the steepest descent method. We note that the solutions to which the conjugate gradient descent method converge are not guaranteed to be global minima. They could be local minima of the functional W. But still, these local minima provide upper bounds on the minimum value of the mix-variance attainable for a fixed value of the action.

4. Examples and numerical results

The examples in this paper are motivated by the experiments performed by Rothstein, Henry & Gollub (1999), where an ordered array of magnets with alternating polarity is used to create an array of vortices with alternating rotations, by means of Lorentz forces in a two-dimensional fluid layer that carries an electric current. In the examples here, we consider two separate arrays with different alignments that can be independently controlled. The kinematics corresponding to this set-up can be captured by the velocity fields (see figure 1)

$$\left.\begin{aligned} \boldsymbol{u}_1(\boldsymbol{x}) &= \begin{bmatrix} -\sin(2\pi x_1)\cos(2\pi x_2) \\ -\cos(2\pi x_1)\sin(2\pi x_2) \end{bmatrix}, \\ \boldsymbol{u}_2(\boldsymbol{x}) &= \begin{bmatrix} -\sin(2\pi(x_1-0.25))\cos(2\pi(x_2-0.25)) \\ \cos(2\pi(x_1-0.25))\sin(2\pi(x_2-0.25)) \end{bmatrix}. \end{aligned}\right\} \tag{4.1}$$

For these velocity fields, the corresponding matrix R as defined in (2.10) can be computed to be $R = \mathrm{diag}\{0.25, 0.25\}$. The initial density field is chosen to be

$$c_0(\boldsymbol{x}) = \sin(2\pi x_2). \tag{4.2}$$

We set $t_f = 1$, $A^* = 1.25$ and make an initial guess of $(\theta)^0(t) = \pi/3$ for the angle function θ. The conjugate gradient descent algorithm described above is performed for 20 iterations. All the computations are done with a uniform grid of 250×250 in the spatial domain and by discretizing the time domain with 250 time steps. A second-order Runge–Kutta method is used for all the Lagrangian particle simulations. Figure 2(*a*) shows the gradient information corresponding to the initial guess and figure 2(*b*) shows the converged solution. Note that the computed optimal solution has

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

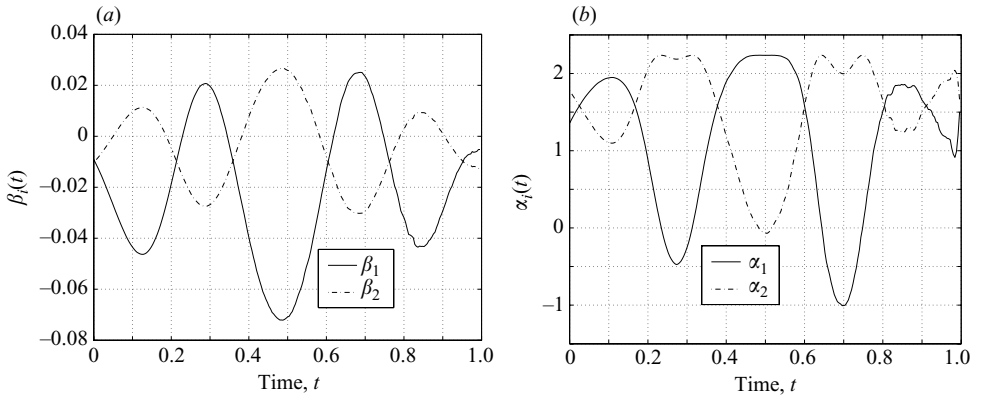


FIGURE 2. (*a*) Gradient information corresponding to an initial guess of $(\theta)^0(t) = \pi/3$. The solid and dash-dot curves are respectively the functions β_1 and β_2 corresponding to the initial guess. (*b*) The optimal solution obtained after 20 iterations of the conjugate gradient descent algorithm. Note that the computed optimal solution has oscillatory components in spite of the initial guess having no oscillations (results for velocity fields defined in (4.1)).

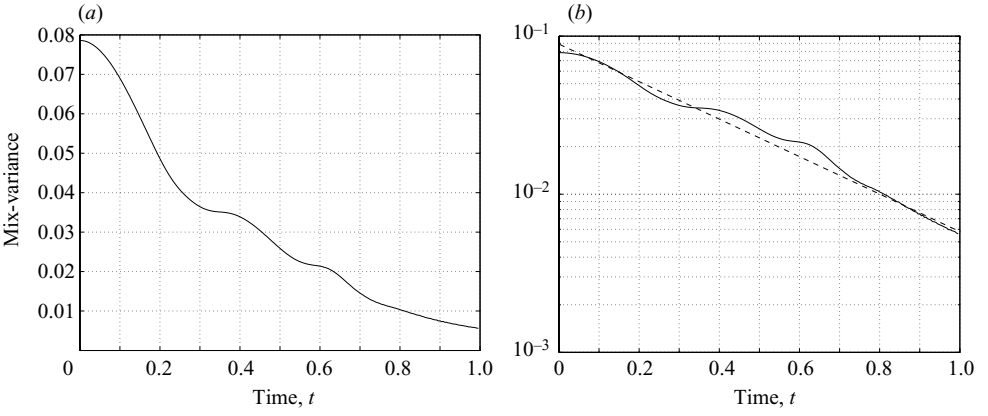


FIGURE 3. (*a*) The decay of the mix-variance with time and (*b*) the corresponding linear-log plot. The decay rate of the mix-variance is not uniform, while the linear–log plot indicates that on average, the mix-variance decays at an exponential rate (results for velocity fields defined in (4.1)).

clear oscillatory components in spite of the initial guess having no oscillations. Figure 3 shows the decay of the mix-variance with time, corresponding to the computed optimal solution. Figure 4 shows snapshots at various times of the evolving density field advected by the optimal control.

In figure 3, it can be observed that the time-derivative of the mix-variance comes close to zero within small time intervals even with the optimally mixing controls. It is possible that no choice of $\boldsymbol{\alpha}(t)$ would make the time-derivative strictly less than zero. For instance, if $c(\boldsymbol{x}, t) = c_s(\boldsymbol{x})$ is an eigenfunction of $[M]$ with eigenvalue s, then

$$\frac{\mathrm{d}\Phi^2(c(.,t))}{\mathrm{d}t} = -2\langle \boldsymbol{u}(.,t)\cdot\nabla c_s(\cdot), [M]c_s(.)\rangle = -2s\langle \boldsymbol{u}(.,t)\cdot\nabla c_s(\cdot), c_s(\cdot)\rangle$$
$$= -s\int_{T^2} \mathrm{div}\left(\boldsymbol{u}c_s^2\right)\mathrm{d}x = 0. \qquad (4.3)$$

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

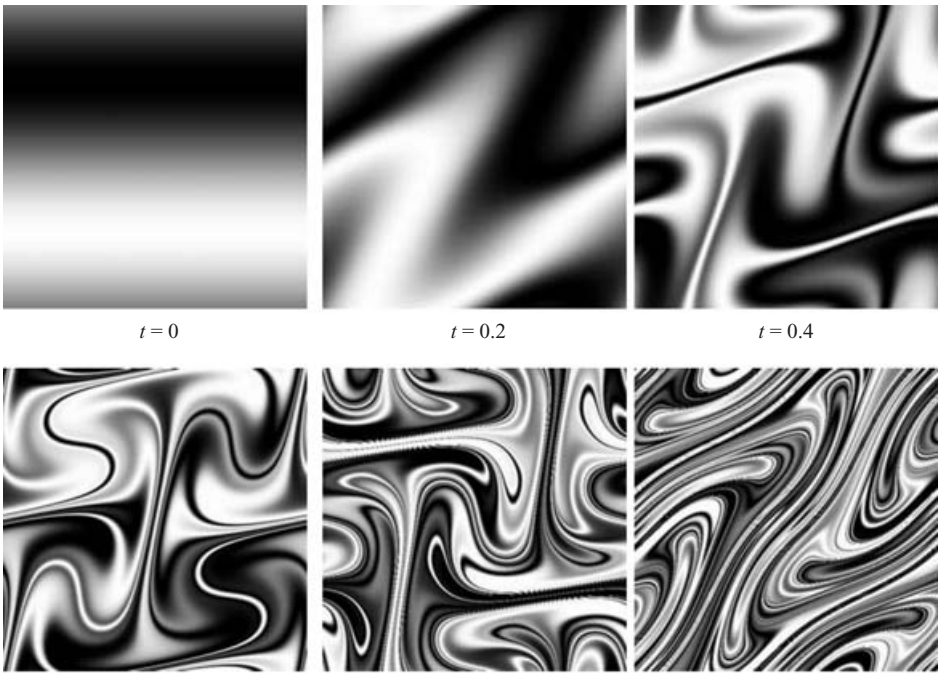


FIGURE 4. Snapshots at various times of the density field advected by the optimal control (results for velocity fields defined in (4.1)).

In fact, for different initial guesses, it is observed that the computed optimal solutions cause a non-monotonic decay of the mix-variance. This stresses the drawback of using a Lyapunov-based feedback method to achieve mixing. A Lyapunov-based feedback method would try to choose values for the controls $\alpha_i(t)$ so as to make the time-derivative of the mix-variance less than zero. But, as shown in the calculation in (4.3), this may be impossible.

As a second example, we use the velocity fields

$$\left.\begin{aligned} \boldsymbol{u}_1(\boldsymbol{x}) &= \begin{bmatrix} -\sin(4\pi x_1)\cos(2\pi x_2) \\ 2\cos(4\pi x_1)\sin(2\pi x_2) \end{bmatrix}, \\ \boldsymbol{u}_2(\boldsymbol{x}) &= \begin{bmatrix} -\sin(4\pi(x_1-0.125))\cos(2\pi(x_2-0.25)) \\ 2\cos(2\pi(x_1-0.125))\sin(2\pi(x_2-0.25)) \end{bmatrix}. \end{aligned}\right\} \tag{4.4}$$

For these velocity fields, the corresponding matrix $\boldsymbol{R}$ can be computed to be $\boldsymbol{R} = \mathrm{diag}\{0.625, 0.625\}$. We use the same set of parameters and initial guess as in the previous example, but set $A^* = 1.0$. Figures 5–8 show the relevant information for this example. Here also, it can be observed that the time-derivative of the mix-variance comes close to zero within small time intervals. Also, oscillatory components appear very clearly in the optimal solutions.

Figure 9 shows how the mixing performance of the optimal controls varies with respect to the action A^*. The computations to generate these plots are done as follows. We start with a low value for the action A^*. We find the optimal solution using the iterative process described above. For a slightly higher value of the action, we use the optimal solution from the previous computation as the initial guess for the iterative process. We repeat this process up to the highest value of action desired.

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

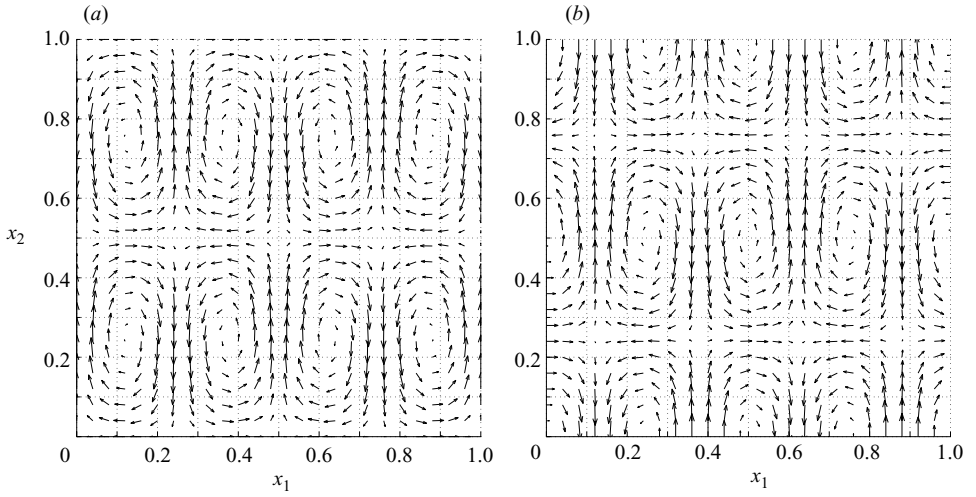


FIGURE 5. Velocity fields defined in (4.4): (*a*) $\boldsymbol{u}_1$, (*b*) $\boldsymbol{u}_2$.

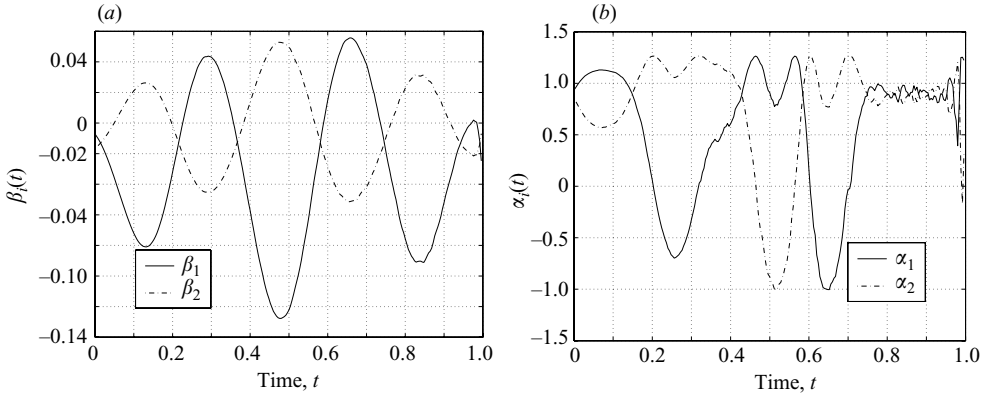


FIGURE 6. (*a*) Gradient information corresponding to an initial guess of $(\theta)^0(t) = \pi/3$. The solid and dash-dot curves are respectively the functions β_1 and β_2 corresponding to the initial guess. (*b*) The optimal solution obtained after 20 iterations of the conjugate gradient descent algorithm (results for velocity fields defined in (4.4)).

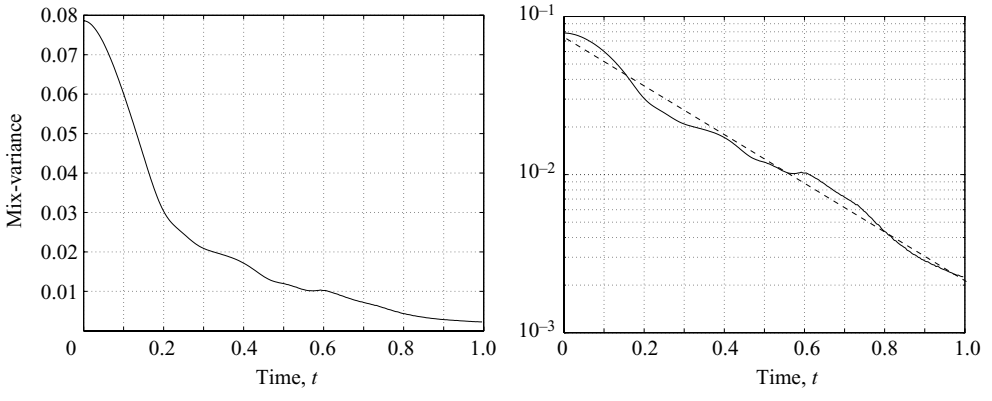


FIGURE 7. (*a*) The decay of the mix-variance with time and (*b*) the corresponding linear-log plot (results for velocity fields defined in (4.4)).

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

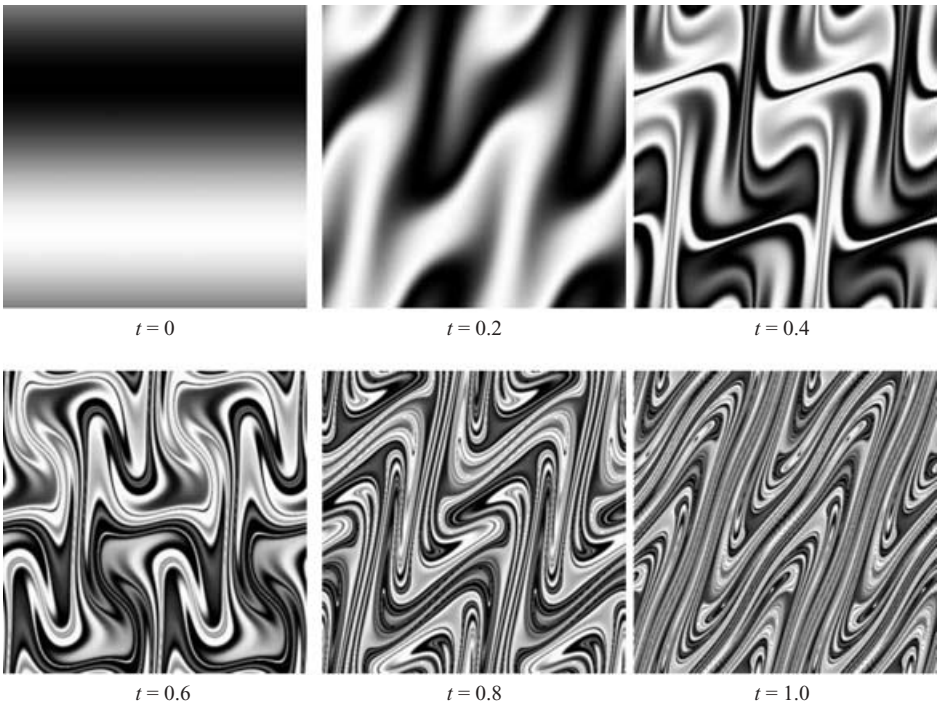


FIGURE 8. Snapshots at various times of the density field advected by the optimal control (results for velocity fields defined in (4.4)).

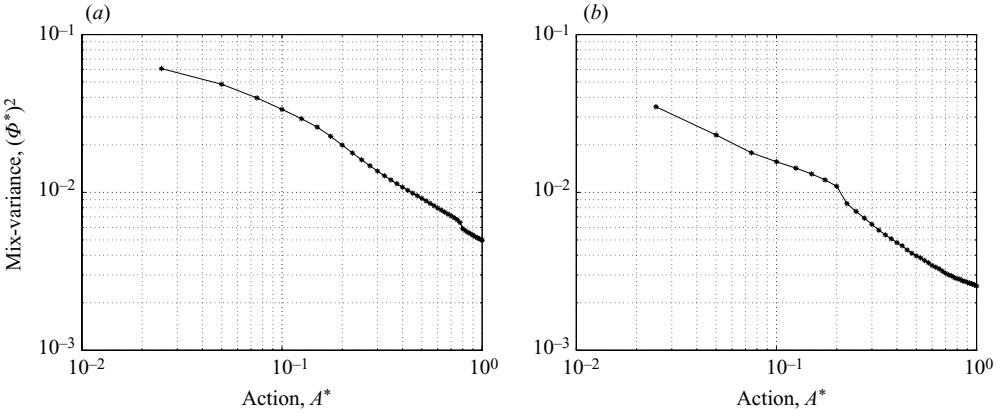


FIGURE 9. The log-log plots for (*a*) velocity fields (4.1) and (*b*) velocity fields (4.4) showing the minimum value of the mix-variance attainable ($(\Phi^*)^2$) as a function of the action A^*.

This makes it likely that for each value of the action, the iterative process starts near to a local minimum and therefore the iteration converges faster. This also makes it more likely that the local minima obtained for two nearby values of the action are not significantly different. From the log–log plots in figure 9, it can be seen that the minimum value of the mix-variance attainable varies almost linearly with respect to the action A^*.

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

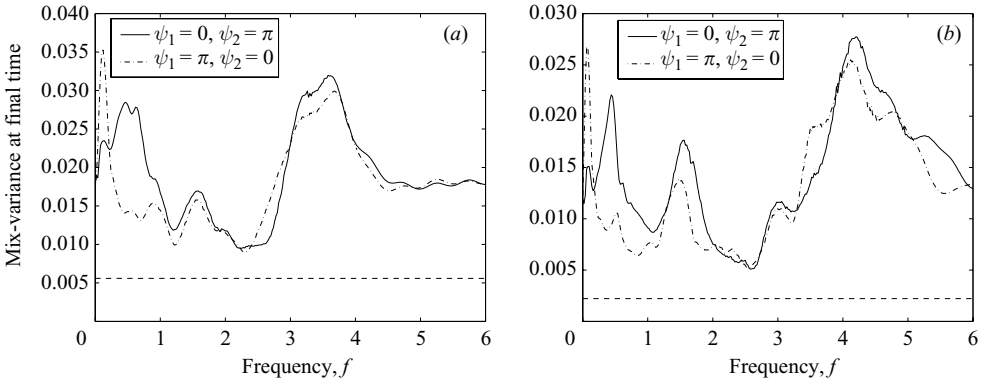


FIGURE 10. Mix-variance at the final time, $\Phi^2(c(.,t_f))$, as a function of the frequency, f, of sinusoidal inputs with mean and L^2 norm comparable to that of the optimal controls α_i^*, corresponding to (*a*) figure 2 (velocity fields (4.1)) and (*b*) figure 6 (velocity fields (4.4)). The horizontal dashed line indicates the level of mixedness achieved by the optimal controls α_i^*.

4.1. *Comparison with sinusoidal inputs*

To compare the mixing effectiveness of the optimal controls $\boldsymbol{\alpha}^*$ with that of sinusoidal inputs, we consider inputs of the form

$$\alpha_i^p(t) = m_i + r_i \sin(2\pi f t + \psi_i). \tag{4.5}$$

So that α_i^p and α_i^* have approximately the same mean and L^2 norm, m_i and r_i are chosen as

$$m_i = \frac{\displaystyle\int_0^{t_f} \alpha_i^*(t)\,\mathrm{d}t}{t_f}, \qquad r_i = \sqrt{\frac{\displaystyle\int_0^{t_f} (\alpha_i^*(t) - m_i)^2\,\mathrm{d}t}{\displaystyle\int_0^{t_f} \sin^2(2\pi t)\,\mathrm{d}t}}. \tag{4.6}$$

Figure 10 compares the mixing performance of sinusoidal inputs with that of the optimal controls computed in the two examples above. The mix-variance at the final time ($\Phi^2(c(.,t_f))$) is plotted as a function of the frequency f for two different cases: ($\psi_1 = 0, \psi_2 = \pi$) and ($\psi_1 = \pi, \psi_2 = 0$). For both examples, it can be seen that the mixing performance of the optimal aperiodic controls is better than that of sinusoidal inputs of all frequencies. The value of the mix-variance achieved by the optimal aperiodic controls is roughly half of that achieved by the sinusoidal input with optimal frequency. This means that the dominant wavenumbers differ roughly by a factor of two.

Figure 11 shows the density fields at the final time when the inputs are sinusoidal with optimal frequencies. A visual inspection of the density fields may not suggest a significant difference in the mixedness achieved by sinusoidal and optimal inputs, but there is a significant quantitative difference as confirmed by the plots in figure 10. We also expect the difference to become more pronounced for higher values of the action. However, at this time, we are unable to demonstrate this numerically. This intuition is confirmed by the results by Pierrehumbert (1994), Rothstein *et al.* (1999) and Liu & Haller (2004). In these works, numerical, experimental and theoretical explorations are made and it is shown that persistent spatial patterns emerge in passive scalar fields when advected by time-periodic flows. These persistent spatial patterns are clearly

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

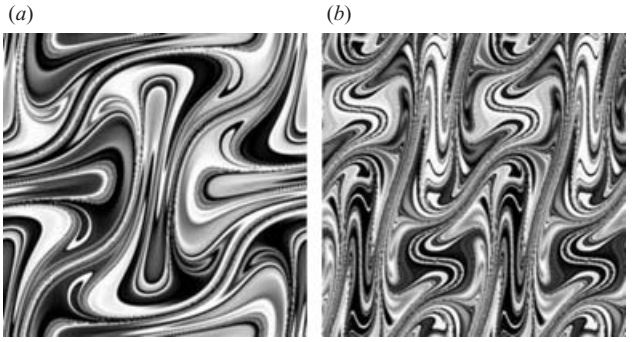


FIGURE 11. Advected density field at the final time $t_f = 1$, when the inputs are sinusoidal, with (a) the same energy as optimal aperiodic controls in figure 2 and with optimal frequency $f_{opt} = 2.25$, and (b) the same energy as optimal aperiodic controls in figure 6 and with optimal frequency $f_{opt} = 2.6$.

not conducive for mixing. However, it must be noted that various flow parameters (frequencies and amplitudes in this case) can be optimized so that time-periodic flows can achieve reasonable mixing, thus possibly making the distinction between periodic inputs and optimal aperiodic controls not very significant for small values of the action.

5. Discussion

Note that the computed optimal solutions for the examples presented in this paper are guaranteed only to be local minima. The solution to which the conjugate gradient descent algorithm converges clearly depends on the initial guess one makes. It may be possible to make finer adjustments to the computed optimal controls to attain better approximations to the optimal solutions. But, given the spatial resolutions with which it is possible to do the computations in a reasonable time, it is not possible to probe the finer scales and make these finer improvements. For our examples, it is observed that there is no substantial change in the cost-function after around 20 iterations of the conjugate gradient descent algorithm.

At present, very little is known about the general structure of the optimally mixing protocols. The infinite-dimensionality of the state makes it difficult to make general comments about the behaviour of the optimal solutions. This is typically the case for high-dimensional optimal control problems. However, we can make one important observation. As discussed in § 3.1, the optimal solutions are such that the vector $\boldsymbol{\alpha}^*(t)$ evolves on the circle for the case when the matrix $\boldsymbol{R}$ is diagonal. Kinematically, this translates into each basis velocity field taking turns in moving material across the streamlines of the other. Still, the exact sequence and timing of switching between the different basis velocity fields has to be numerically computed using methods as described in our paper.

Numerous variants of the optimal control problem posed here can be considered. One version is where the action is not taken into consideration, but the inputs are constrained by inequalities of the form

$$0 \leqslant |\alpha_i(t)| \leqslant \alpha_i^{max}, \tag{5.1}$$

and the objective is to achieve a desired degree of mixedness (say $\Phi^2(c(.,t_f)) = \Phi^2_{des}$) within minimum time ($t_f$). Using Pontryagin's maximum principle, it can be predicted

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

that the optimal controls will be of the *bang-bang* type (see Kirk 1970), i.e. the optimal controls α_i^* would assume only the three values $\{-\alpha_i^{max}, 0, \alpha_i^{max}\}$. The equations describing the evolution of the costate field would be the same as in the TPBVP discussed in this paper, but with a slightly different boundary condition, and the optimal controls would be such that

$$\alpha_i^*(t) = \text{sign}(\langle \lambda^*(.,t), \boldsymbol{u}_i(.)\cdot\nabla c^*(.,t)\rangle)\,\alpha_i^{max}. \tag{5.2}$$

However, the existence of solutions to this problem depends on the issue of controllability. It must be possible to achieve the desired degree of mixedness with the available basis velocity fields. An interesting aspect of this problem is the nature of the switching of the optimal controls between the extreme values. We expect the switching to occur faster and faster with time as the density field becomes more and more mixed. All of these issues will be the subject of future work.

This work was supported by NSF/IGERT DGE-0221715, AFOSR Grant No. F49620-03-1-0096 and NSF DMS-0507256:

Appendix A. Derivation of necessary conditions for optimality and the two-point boundary value problem

The goal is to find the extrema of the functional $W(\boldsymbol{\alpha}) = \langle c(.,t_f), [M]c(.,t_f)\rangle$, where $c(\boldsymbol{x},t)$ is the solution to

$$c_t(\boldsymbol{x},t) + \sum_i \alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c(\boldsymbol{x},t) = 0, \qquad c(\boldsymbol{x},0) = c_0(\boldsymbol{x}), \tag{A 1}$$

and such that

$$\int_0^{t_f} \boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}(t)\,\mathrm{d}t = A^*. \tag{A 2}$$

The partial differential equation (A 1) is treated as a constraint together with the constraint (A 2) and, using the Lagrange multiplier formalism, an augmented functional $\widehat{W}$ is defined as

$$\begin{aligned}\widehat{W}(\boldsymbol{\alpha},c,\boldsymbol{\lambda},z) &= \langle c(.,t_f), [M]c(.,t_f)\rangle - z.\left(A^* - \int_0^{t_f}\boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}(t)\,\mathrm{d}t\right)\\ &\quad - \int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}(\boldsymbol{x},t)\left(c_t(\boldsymbol{x},t) + \sum_i \alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\ &:= W_a(\boldsymbol{\alpha},c,\boldsymbol{\lambda},z) + W_b(\boldsymbol{\alpha},c,\boldsymbol{\lambda},z) + W_c(\boldsymbol{\alpha},c,\boldsymbol{\lambda},z),\end{aligned} \tag{A 3}$$

where

$$\left.\begin{aligned} W_a(\boldsymbol{\alpha},c,\boldsymbol{\lambda},z) &:= \langle c(.,t_f), [M]c(.,t_f)\rangle,\\ W_b(\boldsymbol{\alpha},c,\boldsymbol{\lambda},z) &:= -z.\left(A^* - \int_0^{t_f}\boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}(t)\,\mathrm{d}t\right),\\ W_c(\boldsymbol{\alpha},c,\boldsymbol{\lambda},z) &:= -\int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}(\boldsymbol{x},t)\left(c_t(\boldsymbol{x},t) + \sum_i \alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \end{aligned}\right\} \tag{A 4}$$

The first variations of the individual functionals W_a, W_b and W_c are given as follows:

$$\partial W_a(\partial\boldsymbol{\alpha},\partial c,\partial\boldsymbol{\lambda},z) = 2\,\langle [M]c(.,t_f), \partial c(.,t_f)\rangle\,. \tag{A 5}$$

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

i.e.

$$W_a(\boldsymbol{\alpha}+\boldsymbol{\partial\alpha}, c+\partial c, \boldsymbol{\lambda}+\partial\boldsymbol{\lambda}, z+\partial z) = W_a(\boldsymbol{\alpha}, c, \boldsymbol{\lambda}, z) + \partial W_a(\boldsymbol{\partial\alpha}, \partial c, \partial\boldsymbol{\lambda}, \partial z) + \text{higher order terms.} \quad (A\,6)$$

And,

$$\partial W_b(\boldsymbol{\partial\alpha}, \partial c, \partial\boldsymbol{\lambda}, \partial z) = 2z.\int_0^{t_f} \boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\partial\alpha}(t)\,\mathrm{d}t - \left(A^* - \int_0^{t_f}\boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}(t)\,\mathrm{d}t\right)\partial z. \quad (A\,7)$$

Now computing $W_c(\boldsymbol{\alpha}+\boldsymbol{\partial\alpha}, c+\partial c, \boldsymbol{\lambda}+\partial\boldsymbol{\lambda}, z+\partial z)$, we obtain

$$\begin{aligned}
&-W_c(\boldsymbol{\alpha}+\boldsymbol{\partial\alpha}, c+\partial c, \boldsymbol{\lambda}+\partial\boldsymbol{\lambda}, z+\partial z)\\
&= \int_0^{t_f}\int_{T^2}(\boldsymbol{\lambda}+\partial\boldsymbol{\lambda})\left((c_t+\partial c_t) + \sum_i(\alpha_i+\partial\alpha_i)\boldsymbol{u}_i\cdot\nabla(c+\partial c)\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\
&= \int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}\left((c_t+\partial c_t) + \sum_i(\alpha_i+\partial\alpha_i)\boldsymbol{u}_i\cdot\nabla(c+\partial c)\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\
&\quad + \int_0^{t_f}\int_{T^2}\partial\boldsymbol{\lambda}\left((c_t+\partial c_t) + \sum_i(\alpha_i+\partial\alpha_i)\boldsymbol{u}_i\cdot\nabla(c+\partial c)\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\
&= \int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}\Big(c_t+\partial c_t + \sum_i\alpha_i\boldsymbol{u}_i\cdot\nabla c\\
&\quad + \sum_i\alpha_i\boldsymbol{u}_i\cdot\nabla\partial c + \sum_i\partial\alpha_i\boldsymbol{u}_i\cdot\nabla c\Big)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + \text{higher order terms}\\
&\quad + \int_0^{t_f}\int_{T^2}\partial\boldsymbol{\lambda}\left(c_t + \sum_i\alpha_i\boldsymbol{u}_i\cdot\nabla c\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + \text{higher order terms.} \quad (A\,8)
\end{aligned}$$

Therefore,

$$\begin{aligned}
&W_c(\boldsymbol{\alpha}+\boldsymbol{\partial\alpha}, c+\partial c, \boldsymbol{\lambda}+\partial\boldsymbol{\lambda}, z+\partial z) - W_c(\boldsymbol{\alpha}, c, \boldsymbol{\lambda}, z)\\
&= -\int_0^{t_f}\int_{T^2}\left(\boldsymbol{\lambda}\partial c_t + \boldsymbol{\lambda}\sum_i\alpha_i\boldsymbol{u}_i\cdot\nabla\partial c + \boldsymbol{\lambda}\sum_i\partial\alpha_i\boldsymbol{u}_i\cdot\nabla c\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\
&\quad - \int_0^{t_f}\int_{T^2}\partial\boldsymbol{\lambda}\left(c_t + \sum_i\alpha_i\boldsymbol{u}_i\cdot\nabla c\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + \text{higher order terms}\\
&:= \partial W_c(\boldsymbol{\partial\alpha}, \partial c, \partial\boldsymbol{\lambda}, \partial z) + \text{higher order terms,} \quad (A\,9)
\end{aligned}$$

where

$$\begin{aligned}
\partial W_c(\boldsymbol{\partial\alpha}, \partial c, \partial\boldsymbol{\lambda}, \partial z) :=& -\int_0^{t_f}\int_{T^2}\left(\boldsymbol{\lambda}\partial c_t + \boldsymbol{\lambda}\sum_i\alpha_i\boldsymbol{u}_i\cdot\nabla\partial c + \boldsymbol{\lambda}\sum_i\partial\alpha_i\boldsymbol{u}_i\cdot\nabla c\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\
&- \int_0^{t_f}\int_{T^2}\partial\boldsymbol{\lambda}\left(c_t + \sum_i\alpha_i\boldsymbol{u}_i\cdot\nabla c\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \quad (A\,10)
\end{aligned}$$

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

Now, applying integration by parts to the first two terms on the right-hand side of (A 10), we obtain

$$\int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}(\boldsymbol{x},t)\partial c_t(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t=\int_{T^2}\boldsymbol{\lambda}(\boldsymbol{x},t_f)\partial c(\boldsymbol{x},t_f)\,\mathrm{d}\boldsymbol{x}-\int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}_t(\boldsymbol{x},t)\partial c(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t, \tag{A 11}$$

and

$$\begin{aligned}\int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}\sum_i\alpha_i\boldsymbol{u}_i\cdot\nabla\partial c\mathrm{d}\boldsymbol{x}\,\mathrm{d}t&=\int_0^{t_f}\int_{T^2}\left(\sum_i\boldsymbol{\lambda}\alpha_i\boldsymbol{u}_i\right)\cdot\nabla\partial c\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\&=\int_0^{t_f}\int_{\Gamma}\left(\sum_i\boldsymbol{\lambda}\alpha_i\boldsymbol{u}_i\right)\cdot\eta\partial c\;\mathrm{d}\Gamma\mathrm{d}t-\int_0^{t_f}\int_{T^2}\nabla\cdot\left(\sum_i\boldsymbol{\lambda}\alpha_i\boldsymbol{u}_i\right)\partial c\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\&=\int_0^{t_f}\int_{\Gamma}\left(\sum_i\boldsymbol{\lambda}\alpha_i\boldsymbol{u}_i\right)\cdot\eta\partial c\;\mathrm{d}\Gamma\mathrm{d}t-\int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}\nabla\cdot\left(\sum_i\alpha_i\boldsymbol{u}_i\right)\partial c\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\&\quad-\int_0^{t_f}\int_{T^2}\left(\sum_i\alpha_i\boldsymbol{u}_i\right)\cdot\nabla\boldsymbol{\lambda}\,\partial c\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.\end{aligned} \tag{A 12}$$

In the expressions above, Γ is the surface bounding the domain and η is the normal to the surface bounding the domain. Since the domain in these discussions is a torus, the first term on the right-hand side of (A 12) is zero. On general domains too, the first term is zero if there is no wall-normal velocity. The second term on the right hand-side of (A 12) is zero too because each $\boldsymbol{u}_i$ is divergence free. Therefore, we have

$$\begin{aligned}\partial W_c(\boldsymbol{\partial\alpha},\partial c,\partial\boldsymbol{\lambda},\partial z)=&-\int_{T^2}\boldsymbol{\lambda}(\boldsymbol{x},t_f)\partial c(\boldsymbol{x},t_f)\,\mathrm{d}\boldsymbol{x}+\int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}_t(\boldsymbol{x},t)\partial c(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\&+\int_0^{t_f}\int_{T^2}\left(\sum_i\alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla\boldsymbol{\lambda}(\boldsymbol{x},t)\right)\partial c(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\&-\int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}(\boldsymbol{x},t)\left(\sum_i(\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c(\boldsymbol{x},t))\,\partial\alpha_i(t)\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\&-\int_0^{t_f}\int_{T^2}\left(c_t(\boldsymbol{x},t)+\sum_i\alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c(\boldsymbol{x},t)\right)\partial\boldsymbol{\lambda}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.\end{aligned} \tag{A 13}$$

The necessary conditions for an extremum require that the sum of all first variations of the augmented functional be zero. Thus, requiring $\partial W_a+\partial W_b+\partial W_c=0$, for all $\boldsymbol{\partial\alpha}$, ∂c, $\partial\boldsymbol{\lambda}$ and ∂z, gives the set of equations

$$\left.\begin{aligned}&c_t(\boldsymbol{x},t)+\sum_i\alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c(\boldsymbol{x},t)=0,\qquad c(\boldsymbol{x},0)=c_0(\boldsymbol{x}),\\&\boldsymbol{\lambda}_t(\boldsymbol{x},t)+\sum_i\alpha_i(t)\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla\boldsymbol{\lambda}(\boldsymbol{x},t)=0,\qquad\boldsymbol{\lambda}(\boldsymbol{x},t_f)=2[M]c(\boldsymbol{x},t_f),\\&\int_0^{t_f}\boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\boldsymbol{\alpha}(t)\,\mathrm{d}t=\mathrm{A}^*,\end{aligned}\right\} \tag{A 14}$$

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

and

$$2z.\int_0^{t_f}\boldsymbol{\alpha}(t)\cdot\boldsymbol{R}\partial\boldsymbol{\alpha}(t)\,\mathrm{d}t=\int_0^{t_f}\int_{T^2}\boldsymbol{\lambda}(\boldsymbol{x},t)\left(\sum_i(\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c(\boldsymbol{x},t))\,\partial\alpha_i(t)\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t$$

$$=\int_0^{t_f}\sum_i\left(\int_{T^2}\boldsymbol{\lambda}(\boldsymbol{x},t)\,(\boldsymbol{u}_i(\boldsymbol{x})\cdot\nabla c(\boldsymbol{x},t))\,\partial\alpha_i(t)\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t$$

$$=\int_0^{t_f}\sum_i\langle\boldsymbol{\lambda}(.,t),\boldsymbol{u}_i(.)\cdot\nabla c(.,t)\rangle\,\partial\alpha_i(t)\,\mathrm{d}t$$

$$=\int_0^{t_f}-\boldsymbol{\beta}(t)\cdot\partial\boldsymbol{\alpha}(t)\,\mathrm{d}t,\tag{A 15}$$

where

$$\beta_i(t):=-\langle\boldsymbol{\lambda}(.,t),\boldsymbol{u}_i(.)\cdot\nabla c(.,t)\rangle\ \text{for}\ i=1,2,\ldots,n.\tag{A 16}$$

Therefore

$$2z\,\boldsymbol{R}\boldsymbol{\alpha}(t)+\boldsymbol{\beta}(t)=0.\tag{A 17}$$

Equations (A 14) and (A 17) constitute the two-point boundary value problem in (3.2).

REFERENCES

Aref, H. 1984 Stirring by chaotic advection. *J. Fluid Mech.* **143**, 1–21.

Arnold, V. I. & Avez, A. 1968 *Ergodic Problems of Classical Mechanics*. Benjamin.

Bajer, K. & Moffatt, H. K. 1990 On a class of steady confined Stokes flows with chaotic streamlines. *J. Fluid Mech.* **212**, 337–363.

Ball, J., Marsden, J. E. & Slemrod, M. 1982 Controllability for distributed bilinear systems. *SIAM. J. Control Optim.* **20**, 575.

Balogh, A., Aamo, O. M. & Kristic, M. 2005 Optimal mixing enhancement in 3-d pipe flow. *IEEE Trans. Control Syst. Tech.* **13**, 27–41.

Banks, S. P. 1987 Optimal control of distributed bilinear systems. *Syst. Control Lett.* **9**, 183–189.

Boyland, P. L., Aref, H. & Stremler, M. A. 2000 Topological fluid mechanics of stirring. *J. Fluid Mech.* **403**, 277–304.

Chrisohoides, A. & Sotiropoulos, F. 2003 Experimental visualization of Lagrangian coherent structures in aperiodic flows. *Phys. Fluids* **15** (3), L25.

D'Alessandro, D., Dahleh, M. & Mezić, I. 1999 Control of Mixing in Fluid Flow: A Maximum Entropy Approach. *IEEE Trans. Automatic Control* **44**, 1852–1864.

Fountain, G. O., Khakhar, D. V., Mezić, I. & Ottino, J. M. 2000 Chaotic mixing in a bounded three-dimensional flow. *J. Fluid Mech.* **417**, 265–301.

Franjione, J. G. & Ottino, J. M. 1992 Symmetry concepts for geometric analysis of mixing flows. *Phil. Trans. R. Soc. Lond.* A **338**, 301–323.

Frigo, M. & Johnson, S. G. 1998 FFTW: An adaptive software architecture for the FFT. *Proc. ICASSP* **3**, 1381–1384.

Gelfand, I. M. & Fomin, S. V. 1963 *Calculus of Variations*. Prentice-Hall.

Grigoriev, R. O. 2005 Chaotic mixing in thermocapillary-driven microdroplets. *Phys. Fluids* **17**, 033601.

Haller, G. & Poje, A. C. 1998 Finite time transport in aperiodic flows. *Physica* D **119**, 352–380.

Jurdjevic, V. & Quinn, J. P. 1978 Controllability and stability. *J. Diffl. Equat.* **28**, 381–389.

Kirk, D. E. 1970 *Optimal Control Theory – An Introduction*. Prentice-Hall.

Lasota, A. & Mackey, M. 1994 *Chaos, Fractals and Noise*. Applied Mathematical Sciences, vol. 97. Springer.

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press

Liu, W. & Haller, G. 2004 Strange eigenmodes and decay of variance in the mixing of diffusive tracers. *Physica* D **188**, 1–39.

Luenberger, D. G. 1984 *Linear and Nonlinear Programming*, 2nd edn. Addison-Wesley.

Malhotra, N. & Wiggins, S. 1998 Geometric structures, Lobe dynamics, and Lagrangian transport in flows with aperiodic time-dependence, with applications to Rossby Wave Flow. *J. Nonlinear Sci.* **8**, 401–456.

Mathew, G., Mezić, I. & Petzold, L. 2005 A multiscale measure for mixing. *Physica* D **211**, 23–46.

Mezić, I. & Wiggins, S. 1994 On the integrability and perturbation of three dimensional fluid flows with symmetry. *J. Nonlinear Sci.* **4**, 157–194.

Ottino, J. M. 1989 *The Kinematics of Mixing: Stretching, Chaos and Transport*. Cambridge University Press.

Petersen, K. 1983 *Ergodic Theory*. Cambridge University Press.

Pierrehumbert, R. T. 1994 Tracer microstructure in the large-eddy dominated regime. *Chaos, Solitons Fractals* **4**, 1091–1110.

Poje, A., Haller, G. & Mezić, I. 1999 The geometry and statistics of mixing in aperiodic flows. *Phys. Fluids* **11**, 2963.

Rothstein, D., Henry, E. & Gollub, J. P. 1999 Persistent patterns in transient chaotic fluid mixing. *Nature* **401**, 770–772.

Sharma, A. & Gupte, N. 1997 Control methods for problems of mixing and coherence in chaotic maps and flows. *Pramana – J. Phys.* **48**, 231–248.

Slemrod, M. 1978 Stabilization of bilinear control systems with applications to non-conservative problems in elasticity. *SIAM J. Control* **16**, 131–141.

Stone, H. A., Nadim, A. & Strogatz, S. H. 1991 Chaotic streamlines inside drops immersed in steady Stokes flows. *J. Fluid Mech.* **232**, 629–646.

Stone, H. A., Stroock, A. D. & Ajdari, A. 2004 Engineering flows in small devices: Microfluidics toward a lab-on-a-chip. *Annu. Rev. Fluid Mech.* **36**, 381–411.

Stone, Z. B. & Stone, H. A. 2005 Imaging and quantifying mixing in a model droplet micromixer. *Phys. Fluids* **17**, 063103.

Stroock, A. D., Dertinger, S. K. W., Ajdari, A., Mezić, I. & Stone, H. A. 2002 Chaotic mixer for microchannels. *Science* **295**, 647–651.

Thiffeault, J.-L. & Finn, M. D. 2006 Topology, Braids, and Mixing in Fluids. *Phil. Trans. R. Soc. Lond.* A **364**, 3251–3266.

Wang, Y., Haller, G., Banaszuk, A. & Tadmor, G. 2003 Closed-loop Lagrangian separation control in a bluff body shear flow model. *Phys. Fluids* **15**, 2251–2266.

Wiggins, S. 1992 *Chaotic Transport in Dynamical Systems*. Springer.

https://doi.org/10.1017/S0022112007005332 Published online by Cambridge University Press